PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification: **G06F 15/21**	A1	(11) International Publication Number: **WO 00/02138** (43) International Publication Date: 13 January 2000 (13.01.2000)

(21) International Application Number: PCT/US99/15236

(22) International Filing Date: 06 July 1999 (06.07.1999)

(30) Priority Data:
60/091,750 06 July 1998 (06.07.1998) US

Published

(60) Parent Application or Grant
BIOS GROUP LP [/]; (). BONABEAU, Eric [/]; (). KAUFFMAN, Stuart, A. [/]; (). PALMER, Richard [/]; (). MORRIS, Francis, E. ; ().

(54) Title: **A METHOD FOR PERFORMING MARKET SEGMENTATION AND FOR PREDICTING CONSUMER DEMAND**

(54) Titre: **PROCEDE PERMETTANT D'EFFECTUER LA SEGMENTATION D'UN MARCHE ET DE PREDIRE LA DEMANDE DES CONSOMMATEURS**

(57) Abstract

The present invention presents a method for partitioning that provides both a relevant metric and a set of clusters through an evolutionary learning process. The present invention further presents a method for determining consumer demand (304) that finds the context dependent, or combinatorial optimized set of properties, uses, or customer features that optimize the value of a product to the customer base. The present invention further includes a framework for the marketing and introduction of novel products. The framework has means to model customers and derive an optimal set of goods (308) to produce alone or in the face of a coevolving competitive environment where other firms are introducing and modifying their own goods.

(57) Abrégé

Procédé de division qui fournit à la fois des mesures pertinentes et une série de types de clients par un processus d'apprentissage évoluant. La présente invention concerne en outre un procédé permettant de déterminer la demande (304) des consommateurs, qui trouve la série optimisée, combinatoire ou dépendante du contexte, de propriétés, d'habitudes ou de caractéristiques des consommateurs permettant d'optimiser la valeur d'un produit par rapport à la clientèle. Elle concerne encore un cadre pour le marketing et l'introduction de nouveaux produits. Ledit cadre possède des moyens permettant de modéliser la clientèle et d'en déduire une série optimale de biens (308) à produire seuls ou dans le contexte d'un environnement compétitif évoluant parallèlement dans lequel d'autres sociétés introduisent et modifient leurs propres produits.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G06F 15/21	A1	(11) International Publication Number: WO 00/02138 (43) International Publication Date: 13 January 2000 (13.01.00)

(21) International Application Number: PCT/US99/15236

(22) International Filing Date: 6 July 1999 (06.07.99)

(30) Priority Data:
60/091,750 6 July 1998 (06.07.98) US

(71) Applicant: BIOS GROUP LP [US/US]; 317 Paseo de Peralta, Santa Fe, NM 87501 (US).

(72) Inventors: BONABEAU, Eric; 377 Calle Loma Norte, Santa Fe, NM 87501 (US). KAUFFMAN, Stuart, A.; 15 Montecito Road, Santa Fe, NM 87501 (US). PALMER, Richard; 4030 Forestdale Drive, Durham, NC 27712 (US).

(74) Agents: MORRIS, Francis, E. et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).

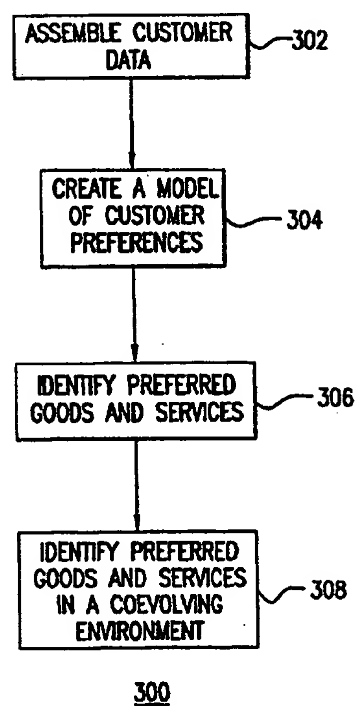
(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: A METHOD FOR PERFORMING MARKET SEGMENTATION AND FOR PREDICTING CONSUMER DEMAND

(57) Abstract

The present invention presents a method for partitioning that provides both a relevant metric and a set of clusters through an evolutionary learning process. The present invention further presents a method for determining consumer demand (304) that finds the context dependent, or combinatorial optimized set of properties, uses, or customer features that optimize the value of a product to the customer base. The present invention further includes a framework for the marketing and introduction of novel products. The framework has means to model customers and derive an optimal set of goods (308) to produce alone or in the face of a coevolving competitive environment where other firms are introducing and modifying their own goods.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Description

A METHOD FOR PERFORMING MARKET SEGMENTATION AND FOR PREDICTING CONSUMER DEMAND

FIELD OF THE INVENTION

The present invention relates generally to methods for performing market segmentation and to methods predicting consumer demand. More specifically, the present invention performs market segmentation by determining dissimilarity measures and predicts consumer demand by constructing a landscape of consumer demand through focused sampling.

BACKGROUND

The problem of determining what combination of factors in a given product (from toothpastes to cookies or cars) will attract customers is a difficult one because the relationship between the factors may be highly nonlinear and because the ratings of factors by customers may be biased for various reasons (formulation of the questions, rating scale, customer answers do not reflect actual preferences, etc.). For marketing purposes, a company would like to be able to find clusters composed of a sufficient number of customers with similar preferences in order to either launch a new product, or adjust an existing product, adapted to such preferences. Of critical importance is that customers with similar sets of preferences be assigned to the same cluster and that customers with significantly different sets of preferences be assigned to different clusters.

Clustering algorithms exist that can generate clusters that satisfy either one or both of these constraints. Multidimensional scaling methods go one step further to allow visualization of high-dimensional data clusters in a low-dimensional embedding space. But clustering algorithms and multidimensional scaling methods always assume the existence of a well-defined metric or dissimilarity measure in attribute space, here the space of factors that contribute to a product.

Accordingly, there exists a need for a method for partitioning that provides both a relevant metric and a set of clusters.

Next, a large body of mathematical and statistical work exists concerning means to estimate the optimal composition of a good or service for a given customer, or population of customers. This body of work contains techniques, known in the art, such as CART, and discrete choice providing means for determining utility functions over a space of properties of a good or service for a given consumer, as well as means of considering a population of different customers with different preferences over that space of properties and attempting to "segment" the customer population into subgroups which may then be specifically targeted by marketing, or shifting the property mix of each product produced and the vector of products to "match" optimally the customer population. Typically the aim is to maximize profit for the firm.

The means in the art, in general, attempt to fit the observed data points by building up sketches of the utility surface for a given consumer or class of consumers using, in the simplest case, linear regression of the data points on all the property axes. Different classes of consumers are discriminated by discovering different linear regression patterns for different, e.g., demographic classes. In more sophisticated approaches, attempts are made to model the possibly curved properties of 'isoultility" surfaces in the space of properties for a given consumer, or class of consumers, by fitting higher order polynomials to the sampled data. The generic problem with this approach is that the data sampled must be used to estimate the coefficients of the monomial and polynomial terms, and the finite amount of data is typically used to characterize the lowest order terms, monomial, quadratic, etc, first. The data is typically "used up in obtaining reliable statistical estimates of these low order terms, and little or none is left over for use estimating higher order terms.

On the other hand, the higher order terms are precisely the measures of the complex context dependent interactions among properties of a single good or service or a collection of goods or services. A trivial example is breakfast, consisting of ham and eggs. These two are traditionally called "consumption complements" by economists. The utility of ham for many consumers is much higher in the presence of eggs, than alone; so too with eggs. Another case is niche marketing of cellular telephones. Not only are these phones of interest to high volume users with expense accounts, but to low volume users who happen to be women with small children driving in rural areas and worried about an accident and no means of calling for help. The context dependence of the properties: woman, mother, with child in car, accident and safety demonstrates the combinatorial character of one niche occupied by this product.

On a simpler level, a given product, say soap, might be characterized by a number of features: color, shape, smell, saponin content, mass, etc. Or coca-cola packaging may be characterized by a number of properties, number of cans, size of cans, fluid in can, color of package, etc.

Accordingly, there is also a need for a method for determining consumer demand that finds the context dependent, or combinatorial optimized set of properties, uses, or customer features that optimize the value of a product to the customer base.

SUMMARY OF THE INVENTION

The present invention presents a method for partitioning that provides both a relevant metric and a set of clusters through an evolutionary learning process.

It is an aspect of the present invention to present a method for partitioning a space of data comprising the steps of:

choosing a plurality of dissimilarity measures;

partitioning the space for each of said plurality of dissimilarity measures;

evaluating said partitioning for each of said plurality of dissimilarity measures; and

selecting one or more of said dissimilarity measures on the basis of said evaluation.

The present invention further presents a method for determining consumer demand that finds the context dependent, or combinatorial optimized set of properties, uses, or customer features that optimize the value of a product to the customer base.

It is an aspect of the present invention to present a method for determining customer demand for products comprising the steps of:

defining a space having R dimensions wherein each point in said space corresponds to a vector of properties;

constructing a landscape for said space comprising the steps of:

locating at least one point on said space where a predetermined customer would purchase a product having said corresponding vector of properties at a predetermined price; and

sampling a set of points on an R-dimensional sphere surrounding said selected point at a predetermined step length from said selected point to determine a first subset of said set of points where the predetermined customer would make a purchase at said predetermined price and to determine a second subset of said sampled points where the customers would not make the purchase at said predetermined price, said first subset of points and said second subset of points form at least one indifference surface between a buying region and a non-buying region at said predetermined price.

The present invention further includes a framework for the marketing and introduction of novel products. The framework has means to model customers and to derive an optimal set of goods to produce alone or in the face of a

coevolving competitive environment where other firms are introducing and modifying their own goods.

It is a further aspect of the present invention to present a method for creating a model of consumer preferences from consumer data comprising the steps of:

constructing a plurality of candidate maps form the consumer data to actual consumer preferences;

constructing a family of agent-based models;

evaluating said plurality of candidate maps and said family of agent-based models with respect to said consumer data;

selecting one or more of said plurality of candidate maps and said family of agent based models based on said evaluation; and

performing at least one operation on said selected candidate maps and said selected agent-based models to generate a new plurality of candidate maps and a new family of agent-based models.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 provides a flow diagram of the adaptive dissimilarity partitioning method 100 of the present invention.

FIG. 2 provides a flow diagram of a method for determining consumer demand 200 that finds the context dependent, or combinatorial optimized set of properties, uses, or customer features that optimize the value of a product to the customer base.

FIG. 3 provides a flow diagram of the framework 300 for the marketing and introduction of novel products.

FIG. 4 provides a flow diagram of the method for creating a model of customer preferences.

FIG. 5 discloses a representative computer system in conjunction with which the embodiments of the present invention may be implemented.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

The present invention presents methods for partitioning that provide both a relevant metric and a set of clusters through an evolutionary learning process called an adaptive dissimilarity partitioning method. Without limitation, many of the following embodiments of the adaptive dissimilarity partitioning method are explained in the illustrative context of market segmentation. However, it will be apparent to persons of ordinary skill in the art that the aspects of the embodiments of the invention are also applicable in any context in which the natural metric or dissimilarity measure of attribute space is not precisely known.

Let us define a set of n m-dimensional data vectors x_i (i=1, ..., n). The components x_{ij} (j=1,..., m) may be real variables, binary variables, or other types of variables. The aim of a typical clustering algorithm is to assign the data points to clusters to minimize some cost function. A prototype vector is usually associated with each cluster: a cluster is defined as the set of data vectors that are closer to the cluster's prototype than to any other prototype. For example, in the k-means clustering algorithm, one has to determine the coordinates of k prototype vectors y_h (h=1, ..., k) to minimize the following cost function:

$$E_k = \sum_{i=1}^{n}\sum_{h=1}^{k} m_{ih}\|x_i - y_h\|^2,$$

where m_{ih}=1 if x_i is assigned to cluster h and m_{ih} = 0 otherwise, and $\|...\|$ is a distance in the space of data vectors. The k-means clustering algorithm is explained in Some methods for classification and analysis of multi variate observations, McQueen, J., *Proc. Fifth Berkeley Symp. on Mathematical Statistics and Probability*, Vol. 1 (Le Cam, L. M. & Neyman, J., eds), University of California Press,

Berkeley, CA, 1965, pp. 281-297. An acceptable clustering solution is given by $\{m_{ih}\}$, where each data vector is assigned to one and only one cluster. In the k-means algorithm, the cluster prototypes are initialized with the first k data vectors. A new data vector x_i, $i>k$, is assigned to the closest prototype vector $y_{h(i)}$. The prototype is adjusted in response to x_i, or, more precisely, is moved closer to x_i:

$$y_{h(i)} \leftarrow y_{h(i)} + \frac{1}{\sum_{u=1}^{i} m_{uh(i)}} (x_i - y_{h(i)}).$$

The total adjustment of the prototype is normalized to the number of vectors that have already been assigned to that prototype. A randomized version of this algorithm, supplemented with topological constraints on prototypes, is the self-organizing map, an unsupervised neural network. Unsupervised neural networks are explained in *The self-Organizing Map*, Kohonen, T., 1990, the contents of which are herein incorporated by reference. We have assumed the existence of a well-defined distance $\|\ldots\|$. Sometimes, only pairwise (or higher-order) relationships among vector components are available. In such cases, the cost function to be minimized is the product of the dissimilarities of data vectors assigned to the same cluster.

Multidimensional scaling (MDS) is used to represent data points in a two-or three-dimensional Euclidian space such that pairwise distances in representation space closely match pairwise dissimilarities as explained in *Multidimensional Scaling*, Cox, T. F. & Cox, M. A. A., Chapman & Hall, London, 1994 ("*Multidimensional Scaling*"), the contents of which are herein incorporated by reference. A clustering algorithm can be applied to the representation vectors. Let y_i be the vector that represents data vector x_i.

Let d_{iu} be the distance between two representation vectors y_i and y_u ($d_{iu} = \|y_i - y_u\|$); and D_{iu} the dissimilarity between x_i and x_u. The cost function (also called stress) to be minimized is typically given by:

$$E = \sum_{i=1}^{n} \sum_{u=1}^{n} w_{iu}(d_{iu} - D_{iu})^2,$$

where the weights w_{iu} are introduced to normalize the absolute values of the disparities D_{iu}. A common choice for w_{iu} is

$$w_{iu} = \frac{1}{D_{iu} \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} D_{\alpha\beta}}.$$

Other definitions of stress and algorithms for minimizing stress are surveyed by *Multidimensional Scaling*.

In both clustering and MDS, the initial dissimilarity measure is assumed to be known. Given the dissimilarity, a clustering algorithm provides clusters whereas MDS provides a low-dimensional representation. The obtained clusters or representations critically depend on the choice of the dissimilarity measure. Such a measure is usually defined on the basis of "intuitive" criteria and relies on the "expertise" of the designer. Defining a dissimilarity measure, however, can in principle be automated. Clustering or scaling data, although it is sometimes used for exploratory data analysis, is usually a first "preprocessing" step in a particular task to be performed (compression, understanding, market segmentation, etc.). The performance of clustering or MDS can therefore be measured not only with respect to the cost function or stress to be minimized but also in connection with the task to be

performed. The appropriate dissimilarity measure can be learned in a supervised manner on a training set, tested on a validation set, and applied to new data. The proposed learning algorithm is a genetic algorithm. Genetic algorithms are described in *Genetic Algorithms in Search, Optimization and Machine Learning*, Goldberg, D. E., Addison-Wesley, Reading, MA.,1989 (*Genetic Algorithms in Search, Optimization and Machine Learning*), the contents of which are herein incorporated by reference.

FIG. 1 provides a flow diagram of the adaptive dissimilarity partitioning method 100 of the present invention. In step 102, the method 100 chooses a family of of distance functions or dissimilarity measures. In step 104, the method 100 randomly generates a population of dissimilarity measures $D^v = \{D^v_m\}$ or distance functions d^v in the chosen family, where v is the index of a given dissimilarity measure in that population. Each "individual" v is encoded into a "genotype".

In step 106, the method 100 performs clustering or multidimensional scaling with a given algorithm for each distance function or dissimilarity measure. In step 108, the method 100 evaluates the performance of clustering or multidimensional scaling and assigns fitness to every dissimilarity measure v. In step 110, the method 100 selects individuals on the basis of fitness. In step 112, the method 100 applies operators to selected individuals and pairs of individuals. Preferably, the operations are genetic operators such as mutation and crossover.

In step 114, the method 100 determines whether the partitioning results are satisfactory with respect to the fitness computed in step 108. If the partitioning results are not satisfactory, control returns to step 106 to perform clustering or multidimensional scaling for each new distance functions or dissimilarity measure created in steps 110 and 112. If the partitioning results are satisfactory, control proceeds to step 116 where the method 100 terminates.

The distance function or dissimilarity measure can be represented by a true function of the vectors' coordinates or by a set of pairwise relationships. When only pairwise relationships between data vectors are available one needs to generalize the dissimilarity measure to data vectors which have not been presented. The simplest generalization procedure is to use a locally linear interpolation, using the k nearest neighbors: the dissimilarity between the new vector V and any other vector O is given by the average dissimilarity between the k nearest neighbors of V and O.

The following examples illustrates the operation of the adaptive dissimilarity partitioning method 100. Let us assume for definiteness that each data vector x_i is two-dimensional. The two components of x_i represent two properties of a cookie, for example, sweetness and chewiyness. A set of n customers is asked to determine the respective levels of sweetness and chewyness they like in a cookie, on a scale of 1 to 10 for each property. In addition, each customer is asked to tell which type of cookie he or she is currently using. Assume that k different types of cookies are represented. The distance function in the space of customer preferences is unknown. For example, one factor may be more important than another. A simple family of distance functions is:

$$d_{iu} = \left| f_1(x_{i1}, x_{i2}, x_{u1}, x_{u2})(x_{i1} - x_{u1})^2 + f_2(x_{i1}, x_{i2}, x_{u1}, x_{u2})(x_{i2} - x_{u2})^2 \right|^{1/2},$$

where f_1 and f_2 are, for example, second-degree polynomial functions of their variables. Each function is characterized by 15 parameters, the coefficients of the polynomials. The variations of these parameters is assumed to be restricted to [-10,10]. A clustering algorithm, such as k-means, is applied to the data set using this distance function. The fitness of a distance d^v is given by

$$F^v = \frac{1}{1 + M_{in} + M_{out}},$$

where M_{in} is the number of customers assigned to the same cluster that do no buy the same cookie type and M_{out} is the number of customers assigned to different clusters that buy the same cookie type. Depending on the task at hand, these two types of mismatch can be given different weights.

The best individuals obtained after, say, 1000 generations of the genetic algorithm, correspond to distance functions that allow to obtain the right clusters of customers in the sense described above.

The adaptive dissimilarity partitioning method 100 of the present invention finds the natural dissimilarity measure or distance function in a space of attributes. This function may be unknown. Instead of resorting to ad hoc functions, the method systematically generates a distance function adapted to the task at hand. The obtained distance function reflects the true structure of the space of attributes and therefore can be used, in the context of market segmentation, to cluster customers, extract the "natural" clusters in the data using a non parametric clustering algorithm (that is, one in which in the number of clusters is not predefined), extract the effective dimension of the space of preferences, test product differentiation, improve positioning by product adjustment, and test potential new products, taking into account the cost of moving from one product to another or of launching a new product.

Other significant areas of application include protein data visualization, protein function and structure prediction, dimensionality reduction for virus data sets, general classification and pattern recognition problems, and data visualization, including database visualization and navigation.

In another example, two hundred two-dimensional data vectors were randomly generated. Let x_{i1} and x_{i2} be the x- and y-coordinates of the ith data vector. x_{i1} and x_{i2} are

drawn from a uniform random distribution on [0,1]. Let us assume that x_{i1} and x_{i2} represent customer preferences for two selected features of a given product type, that two products are on the market, and that a customer i purchases product 1 if and only if $x_{i1}<0.5$ and purchases product 2 if and only if $x_{i1} \geq 0.5$. In this example, therefore, only x_{i1} is relevant in the determination of what product is purchased by a customer whose preference vector is (x_{i1}, x_{i2}). But this information is not known to the analyst, who simply assumes that the relevant distance in preference space is, for example, the Euclidian distance. Using such a distance, the analyst will be unable to correctly segregate customers into two classes. What the algorithm has to find is the relevant distance in preference space that will naturally lead to the correct segregation after application of a simple clustering algorithm. Here we use a modified version of the k-means clustering algorithm with k=2. Two centroids are initially located at (0.5, 0.25) and (0.5, 0.75). Ideally, after application of the clustering algorithm with the appropriate distance function, the centroids should converge to (0.25, 0.5) and (0.75, 0.5). Remember that with this clustering algorithm a data vector belongs to the cluster whose centroid is closest to that data vector. Let $C_{m(i)}$ be the centroid closest to vector x_i ($C_{m(i)} = ArgMin\left(d\left(C_m, x_i\right)\right)$, where d is the distance function), and $C_{m(i)j}$ the jth coordinate (j=1,2) of $C_{m(i)}$. The centroid update upon presentation of x_i is given by:

$$C_{m(i)j} \leftarrow C_{m(i)j} + \eta \frac{d(C_{m(i)}, x_i)}{n} \sigma(x_{ij} - C_{m(i)j}),$$

where d is the current distance function, $\sigma()$ is the sign function ($\sigma(u)=+1$ if u>0, $\sigma(u)=-1$ if u<0, and $\sigma=0$ if u=0), η is a learning rate, and n=200 is the number of data vectors.

The family of distance function used in this example has three parameters:

$$d(x_i, x_h) = \left[w|x_{i1} - x_{j1}|^{\alpha} + (2-w)|x_{i2} = x_{j2}|^{\beta} \right]^{\frac{2}{\alpha+\beta}},$$

where w, α, and $\beta \in [0,2]$. When w=1 and α= β=2, the usual Euclidian distance is recovered, and when w=1 and α=β=1 one gets the city-block (or L_1) distance.

This family of distance functions can easily be generalized to higher-dimensional spaces. For example, let us consider a D-dimensional space:

$$d(x_i, x_h) = \left[\sum_{p=1}^{D} w_p |x_{ip} - x_{jp}|^{\alpha_p} \right]^{\frac{D}{\sum_{p=1}^{D} \alpha_p}},$$

with

$$\sum_{p=1}^{D} w_p = D,$$

where α_p(p=1,.. .,D) and w_p (p=1,. .,D) are 2D parameters (of which only 2D-1 are free parameters) that determine the relative importance of the pth coordinate and the amount of distortion along the pth coordinate. This family of functions assumes no correlation among coordinates, which is certainly a limitation in certain cases. Other distance functions should be used in such cases.

For the simple two-dimensional example, a simple, fitness-proportionate (+elitism) genetic algorithm (GA) was used with the following fitness function for distance d^Y:

$$F^v = \frac{1}{1 + M_{in} + M_{out}},$$

where M_{in} is the number of customers assigned to the same cluster that do no purchase the same cookie type and M_{out} is the number of customers assigned to different clusters that buy the same product. The population size was 40, the mutation rate 0.1, and crossover was replaced with averaging of parameters (that is, two selected individuals produce one offspring the parameters of which are the arithmetic average of its parents' parameters). After 10 generations, the GA finds values of the parameters that consistently produce a perfect clustering of customers after application of the modified k-means algorithm. During one application (200 iterations) of the k-means algorithm for "bad" values of the parameters (w=0.96, α=1.81, β=1.77), close to the Euclidian distance: the centroids are unable to move to the optimal locations and remain confined in the vicinity of their initial values. For "good" values of the parameters found by the GA after 10 generations (w=1.98, α=1.67, β=0.03), the centroids move to the optimal locations because the distance function assigns almost all the weight to the x-coordinate. The GA has therefore been able to find a distance function, within the family of distance functions, that reflects the "true" structure of preference space.

Assume now that instead of being uniformly distributed in [0,1] x [0,1] customers form four clusters (with the same "purchase" rule: customer i purchases product 1 if and only if x_{i1}<0.5 and purchases product 2 if and only if x_{i1} ≥0.5). Two situations can occur: the four clusters may discriminate along the y-axis or along the x-axis. Upon application of a non-parametric (number of clusters undefined) clustering or multidimensional scaling algorithm, the situation where the four clusters may discriminate along the y-axis should lead to the detection of 2 clusters while

the situation where the four clusters discriminate along the x-axis should lead to the discovery of 4 clusters if the appropriate distance function is used. If the Euclidian distance function is used both situations lead to the detection of 4 clusters. A non-parametric (ant-based) algorithm leads to 4 clusters in both cases using the Euclidian distance. The same algorithm leads to 2 clusters when applied to the situation where the four clusters discriminate along the y-axis and 4 clusters in the situation where the four clusters discriminate along the x-axis.

In an alternate embodiment, for more complicated problems, general function approximators such as neural networks are used instead of family of distance functions. In the case of neural networks connections weights are evolved using the genetic algorithm.

In another alternate embodiment, GA is interactive: the outcome of the clustering or MDS algorithm is evaluated by a human observer who picks the good solutions.

The adaptive dissimilarity method 100 is also applicable to graph partitioning. Let $G=(V,E)$ be a non-directed graph. $V=\{v_i\}_{i=1,\ldots,n}$ is the set of n vertices and E, a subset of VxV, the set of edges, of cardinal $|E|$. E can be represented as a matrix $[e_{ij}]$ of edge weights, e_{ij} being the weight of edge (v_i,v_j), where $e_{ij} \neq 0$ if $(v_i,v_j) \in E$ and $e_{ij}=0$ if $(v_i,v_j) \notin E$. The bipartitioning problem consists of finding 2 sets of n/2 vertices each such that the total edge weight between clusters is mimimal. This problem is known to be NP-complete, and many heuristics have been proposed to find reasonably good solutions in polynomial time. The question we may ask is the following: is there a natural distance in connection space (where the coordinate of a vertex v_i is given by e_{ij}, $j=1,\ldots, n$) such that the application of the k-means clustering algorithm (k=2) generates a good solution of the bipartitioning problem?

The adaptive dissimilarity partitioning method 100 has been tested on random graphs $G(n, c, p_i, P_e)$, where n=100 is the total number of vertices, c=2 is the number of

clusters, p_1 is the probability that two vertices within a cluster are connected, and p_2 is the probability that two vertices belonging to different clusters are connected. Edges are characterized by $e_{ij}=1$ if $(v_i,v_j) \in E$ and $e_{ij}=0$ if $(v_i,v_j) \in E$. Such graphs are convenient to test the algorithm because the optimal solution of bipartitioning is known when c=2: the optimal partitioning solution consists of having as many vertices as possible that belong to the same graph cluster allocated to the same partition cluster. These graphs have been introduced and are used as difficult benchmark problems in the context of VLSI design.

A modified version of the k-means algorithm is applied. The two centroids are initially assigned random coordinates. Let $C_{n(i)}$ be the centroid closest to vertex v_i ($C_{n(i)} = \operatorname{ArgMin}_{s=1,2}(d(C_s, v_i))$), where d is the current distance function), and $C_{n(i)j}$ the jth coordinate ($j=1,\ldots,n$) of $C_{n(i)}$. The centroid update upon presentation of x_i is given by:

$$C_{n(i)j} \leftarrow C_{n(i)j} + \eta \frac{d(C_{n(i)}, v_i)}{n} \sigma\left(e_{ij} - C_{n(i)j}\right),$$

where d is the current distance function, $\sigma()$ is the sign function ($\sigma(u)=+1$ if $u>0$, $\sigma(u)=-1$ if $u<0$, and $\sigma=0$ if $u=0$), η is a learning rate, and $n=200$ is the number of data vectors. The family of distance function used in this example has three parameters:

$$d(v_l, v_h) = w\left[\sum_u \left|e_{lu} - e_{hu}\right|^{\alpha}\right]^{\frac{1}{\alpha}}$$

$$+(1-w)\left\{\sum_u e_{lu}\sum_l \left|e_{ul} - e_{hl}\right|^{\beta} + \sum_u e_{hu}\sum_l \left|e_{ul} - e_{il}\right|^{\beta}\right\}^{\frac{1}{\beta}}$$

where $w \in [0,1]$, and α and $\beta \in [0,2]$. When $w=1$, one gets usual distances. The first term contains only zeroth and first-order relationships between the two vertices: this term is small when the two vertices are connected (0th-order) and are connected to the same set of vertices (first-order). The second term, which gets activated when $w<1$, represents second-order relationships between two vertices: this term is small when the neighbors of the two vertices have a lot of adjacent vertices in common. Such relationships may be important for graph partitioning, but the extent to which they improve the partitioning is not known.

The fitness function used in the GA for distance d^v is given by:

$$F^v = \frac{1}{1 + E_{int} + \left|n_1 - \frac{n}{2}\right|}$$

where E_{int} is the total inter-cluster weight, and n_1 is the number of vertices assigned to cluster 1. The $\left|n_1 - \frac{n}{2}\right|$ term is there to favor well-balanced solutions.

The present invention further presents a method for

determining consumer demand that finds the context dependent, or combinatorial optimized set of properties, uses, or customer features that optimize the value of a product to the customer base.

Previous work has developed a general model of rugged fitness landscapes called the NK model as explained in *The Origins of Order*, Stuart A. Kauffman, Oxford University Press, 1993, Chapter 2, the contents of which are herein incorporated by reference. The NK model is also explained in *At Home in the Universe*, Stuart A. Kauffman, Oxford University Press, 1995, Chapter 9, the contents of which are herein incorporated by reference.

NK landscapes are members of a still more general class of models in physics, and known in the art as P spin models. A P spin model consists of N spins, each of which can take on a discrete number of values, say -1 and +1, or 1 and 0, or a,b,c,d. Each spin contributes an "energy" to the total energy of a system of N spins. The energy of a given spin configuration of the N spins is given by the sum of the energies of the N spins. Each spin's energy contribution is, in general, given by a sum of a monomial term which is a function of its own state, plus quadratic terms which are sums of energies that are functions of the states of all spins that influence it in pairwise interactions, plus a similar sum of cubic terms listing all the contributions of all triples of spins of which that spin is a member, plus higher order terms. In the NK model, K is the highest order coupling.

In such spin-glass models, the discrete system has a rugged "fitness" "cost" "efficiency" or "utility" landscape over the combinations of states of the N spins. New techniques have been developed to characterize a number of features of such landscapes. And it is these features that allow ready assessment of the importance of higher order, combinatorial properties on landscape structure. These properties include: 1) The number of peaks in the landscape; 2) The expected number of steps to a peak from any given

point in the landscape. 3) The dwindling number of directions "uphill" as the peak is climbed. 4) The number of different peaks that can be climbed from a single point on the landscape by adaptive walks which must proceed only uphill. 5) The correlation structure of the landscape which is, roughly, the correlation between fitnesses at two points on the landscape as a function of their distance.

These properties of discrete landscapes, where the spins take on only discrete values, a,b,c,d... can be generalized to the case of continuous dimensions, where each variable is a real number. This continuous case, the lengths of walks uphill, and dwindling directions uphill must be parameterized by a "step length" in the space of reasonably smooth hill sides, any point on the landscape that is on a hillside has the property that, for infinitesimal steps away from that point, half the directions are uphill and half are downhill. Only on ridges, saddles and peaks is that false. However, if a discrete step length, say 100 yards, is specified, then as a walk continues uphill and a ridge or saddle or peak is approached, the "cone" that is still uphill will dwindle. The rate of dwindling is a measure that can be used to characterize the ruggedness of a continuous landscape. Thus, on NK landscapes, with K modestly large, the generic feature is that at every step uphill, the number of directions uphill falls by a constant fraction. As landscape ruggedness increases, the fraction by which the directions uphill dwindles increases from a few percent to 50% for fully random landscapes in the K - N 1 "random energy" limit. In a similar way, the rate at which the uphill cone decreases as walks uphill continue provides a measure of landscape ruggedness for continuous landscapes.

Consider a product space, without loss of generality taken to be soap. Features of this product were noted above, and in general, include other features of interest. Consider, to be, concrete and without loss of generality, discrete choice methods. A customer is presented with different choices of a bundle of properties, or vector

of properties. Each bundle is a point in the property space. A price is attached to each such point. The customer is asked to choose which, if any, he/she would buy. Examination of the vector is the property space after a finite number of such choices, reveals a price in the vicinity of those positions in property space at which the customer will just stop buying. Thus, on one side of a point on a surface in property space, at that price, the customer will not buy, on the other side of a point on a surface in property space, at that price, the customer will not buy, on the other side he will buy. The point in question estimates the price for that specific vector of properties. By sampling at many points for one customer, it is possible to build up a set of points that estimates the utility curve, or surface, in property space at one price for that customer, hence an indifference surface, and a set of such surfaces at different prices.

For a population of customers, a population of such data points can be assembled. In principle, much data could be obtained from each customer, but typically it is only feasible to obtain a limited amount of data from a given customer. Typically, this is obtained over a moderate large region of property space. The data points are then typically each labeled by a vector of demographic traits, and an attempt is made using standard analysis to discriminate both high utility positions in the space of properties, and simultaneously the targeted demographic populations that are well matched to good positions in the space of properties in order to optimize the vector of goods produced, each at a different position in the property space, and targeted to one or more positions in the demographic space, such that a total figure of merit such as total profit after total manufacture and sales.

The application of landscape ideas can improve these standards procedures both by directing the limited sampling that can be done that it helps capture higher order terms, or context dependent features, of these marketscapes, helps build statistical models of the right "equivalence

class" of the real market scape, and helps build actual models of the actual marketscape.

FIG. 2 provides a flow diagram of a method for determining consumer demand 200 that finds the context dependent, or combinatorial optimized set of properties, uses, or customer features that optimize the value of a product to the customer base. In step 202, the method for determining consumer demand 200 selects a point in property space that lies on a surface that divides a region where a predetermined customer would buy from a region where the predetermined customer would not buy.

In step 204, the method for predicting consumer demand 200 samples a set of points on an R-dimensional sphere surrounding the point selected in step 202. Step 204 contrasts with previous methods for predicting consumer demand that sample widely over the product space. The radius of the sphere is defined in a well specified way where the radius is defined as the "step length" on the surface. An exemplary distance is the Euclidian distance. With the same customer, or more generally, the same class of customers, step 204 characterizes for many points in the spherical surface surrounding the point whose price has been determined, whether that new point would or would not be purchased by the customer at the given price. Since the true price surface in the space of properties contains the first determined point, that price surface will, in general, pierce the spherical surface surrounding the point whose price is determined. The points on the sphere which are purchased and the points which are not purchased determine, in the simplest case, a curve of points whose price is the transition between buying and not buying at the price. In this way, the neighborhood surrounding that first priced point can be examined.

In step 206, the method for predicting consumer demand 200 determines whether the indifference surface has been substantially completed. If the method for predicting consumer demand 200 determines that the indifference surface

has not been substantially completed, control proceeds to step 208. In step 208, the method for predicting consumer demand selects another point on the indifference surface from the transition curve determined in step 204. After step 208, control returns to step 204. Step 204 samples a set of points on an R-dimensional sphere surrounding the point selected in step 208. In this fashion, the method for predicting consumer demand 200 operates to extend the indifference surface at the predetermined price in any direction in the property space.

The ruggedness of the indifference surface at a given price will show up by any of the properties we have discussed. Thus, measured in property space, the indifference surface at a given price may have one or more correlation lengths in the space of properties. These correlation lengths, in the NK model are long, for K small, and short for K large. Thus, short correlation lengths estimate higher order couplings among the properties. The cone "uphill" in property space on an indifference surface at a given price can be determined. Good combinations of properties will show up as peaks or minima, depending upon direction of definition, in the surface. That is, a very good combination of properties in property space will show up, for example, as a willingness to pay the fixed price for a small "amount" of the given vector of properties. Having defined a local "peak" in the indifference landscape surface, we can define the typical walk length, given step size, the peak, and the number of peaks to which one can walk from any point. In addition, we can examine the similarity of peaks climbed from the same or nearby points on the indifference landscape at a given price. We can ask if high peaks cluster near one another. We can ask whether recombination is a good means to find the high peaks. If so, we can search out the high peaks by focusing our questioning in precise ways, to look "between" the high peaks on the current landscape, and hill climb from those points to still higher peaks.

All these properties allow focused sampling of the landscape to estimate the higher order context dependent, combinatorial features of a given market scape.

Statistical models of the sampled market scape can be built by utilizing P spin-like models, where the class of models with all possible values of the coefficients of all the Padic terms in the polynomials constitute the family of landscape models. Maximum entropy Bayesian updating similar techniques can then be used to estimate the most likely landscape parameters to fit the observed data. A major difference between the current approach and usual approaches is that the detailed sampling in specific regions of the indifference surface at a given price yields estimates of the how "high" the higher order terms, (K in the Nk model) actually are. Thus, we can estimate from such focused local measurements at several points on the landscape, that, for example, fifth order interactions, P=5, are critical for determining the local structure of the marketscape. Knowing that, we can use a preponderance of the data to fit or estimate the 5th order terms, and only a small amount of the data to estimate the monomial terms that may determine the overall non-isotropic features of the marketscape on long length scales across the marketscape. Thus, we can optimize use of the sampled data to discover both long range features of the landscape and local features.

Given this analysis, one can derive a class of statistical models of the landscape, and specific models of the landscape.

The method for predicting consumer demand 200 was explained in the context of computing an indifference surface for a predetermined price in the property space for a predetermined customer. However, as is known by one of ordinary skill in the art, the method for predicting consumer demand 200 could also be used to sample the property space of the product for a given class of customer at a predetermined price or at a set of predetermined prices. Further, the method for predicting consumer demand 200 could also be used

to arrange the demographically characterized population of customers into a customer-scape for any given point in the product space. This new approach to market segmentation arises by casting the agents into an M dimensional demographic space. At any given price, we can determine the fraction of customers in any small volume of demographic space who will buy the good at that point in product space at the given price. This determines a "customer-scape" for that good at that price. Once again, the customer-scape is a landscape, and we can define all the properties noticed above: correlation structure, lengths of fixed step length walks to peaks, the dwindling cone uphill as peaks are climbed, the number of peaks accessible from a given point, the similarity of such peaks, and whether high peaks cluster near one another. In the latter case, recombination is a good means to search the landscape. This procedure defines one or more optimal customer features for a given good, or position in product space. The same procedure allows multiple points in product space to be utilized, indeed just the points normally utilized, to find the best set of positions in product space to match the best targeted populations of customers in customer space. Again, the advantage of our procedure is that it allows the higher order terms, the context dependent features in customer space, to be more readily detected, for it tells us that K order terms are important. Again, we can then construct statistical models of customer-scapes, and models of specific customer scapes.

The present invention further includes a framework for the marketing and introduction of novel products, which is a central function of businesses. FIG. 3 provides a flow diagram of the framework 300 for the marketing and introduction of novel products. The framework 300 concerns means to model customers and derive an optimal set of goods to produce alone or in the face of a coevolving competitive environment where other firms are introducing and modifying their own goods.

In step 302, the framework for the marketing and introduction of novel products 300 assembles data on customers from statements of preferences on questionnaires, point of purchase data, neilson data, etc. In step 304, the framework for the marketing and introduction of novel products 300 creates a model of customer preferences. In step 306, the framework 300 uses the models of customer preferences created in step 304 to identify preferred goods and services. In step 308, the framework considers the behavior of other firms in the environment in addition to the models of customer preferences created in step 304 to identify preferred goods and services in a coevolving competitive environment.

FIG. 4 provides a flow diagram of the method for creating a model of customer preferences of step 304. In step 402, the method for creating a model of customer preferences 304 determines whether to perform market segmentation. If step 402 indicates that market segmentation should be performed, control proceeds to step 404 where the method for creating a model of customer preferences executes the adaptive dissimilarity partitioning method 100 shown by the flow diagram of FIG. 1. If step 402 indicates that market segmentation should not be performed, control proceeds to step 406.

In step 406, the method for creating a model of customer preferences 304 constructs a family of linear or non-linear models of customers. These models are candidate maps from answers to questions, point of purchase data, etc. to the actual predictive preferences of the customers for the goods in question. Accordingly, an aim of the method for creating a model of customer preferences 304 is to order the goods in a match to actual preferences of customers.

In step 408, the method for creating a model of customer preferences 304 constructs agent based models of customers based on default hierarchies, rules of thumb, etc. in their strategy space. Default hierarchies, etc. do not require that preferences be transitive, which is often true

of customers. In contrast, a preference space does require transitivity. Agent based models of customers are described in *A System and Method for the Synthesis of an Economic Web and the Identification of New Market Niches*, Attorney docket number 9392-0007-999, filed May 15, 1998, the contents of which are herein incorporated by reference. Agent based models of customers are further described in *An Adaptive and Reliable System and Method for Operations Management*, Attorney docket number 9392-0004-999, filed July 1, 1999, the contents of which are herein incorporated by reference.

In step 410, the method for creating a model of customer preferences 304 utilizes adaptive algorithms over the space of mappings produced by step 406 and the space of agent strategies produced by step 408 to find a set of models that predicts customer purchasing preferences for a set of goods. In the preferred embodiment, the adaptive algorithms are genetic algorithms. In an alternate embodiment, the adaptive algorithms are genetic programming.

In step 412, the method for creating a model of customer preferences 304 determines whether the output of step 410 has produced good predictive models of customer purchasing preferences. If step 412 determines that the output of step 410 has not produced good predictive models of customer purchasing preferences, control returns to step 406 where processing proceeds with the new set of models produced by the adaptive algorithm of step 410. If step 412 determines that the output of step 410 has produced good predictive models of customer purchasing preferences, control proceeds to step 414 where the processing terminates.

As previously discussed, in step 306, the framework 300 uses the models of customer preferences created in step 304 to identify preferred goods and services. If the customers have preferences for may features of a product that add up to a single preference landscape, then step 306 executes the method for predicting consumer demand 200 illustrated by the flow diagram of FIG. 2. In contrast, if the customer preferences are not commensurable, then step 306

executes an optimization tool to find the global pareto optimal points such as Configuration Sherpa, which is described in *A System and Method for Coordinating Economic Activities Within and Between Economic Agents*. In either case, one of ordinary skill in the art would understand that there are a variety of clustering and multi-dimensional scaling algorithms that can seek optimal choices of locations of goods in the product space to attract the most customers. Such algorithms may prespecify the number of goods, or seek optimal numbers and locations of goods based on a firm's budget constraints, and other aspects of firm operations in its competitive environment.

As previously explained, in step 308, the framework considers the behavior of other firms in the environment in addition to the models of customer preferences created in step 304 to identify preferred goods and services in a coevolving competitive environment. Firms compete by introducing or improving products. Hence, there is a coevolutionary dynamic. Generically, there are two regimes: a "red queen" regime of persistent coevolution in the space of products and an evolutionary stable strategies regime where all products reach local or global Nash equilibria and stop moving in product space. See *At Home in the Universe*. If the firm completes the observe, orient, decide and act loop (OODA) faster than the other firms with respect to the introduction, innovation, improvement and wise placement of products, it can systematically win.

Step 308 of the framework for the marketing and introduction of novel products 300 uses models of customers and capacity to predict preferences over the space of products to build agent based or other dynamical models of the coevolution of market shares of products, utilizing data to locate optimal positions for new or improved products in coevolutionary dynamics subject to constraints on budget, capacity, and time to market for new or improved goods, etc.

Agent based models that identify new products are described in *A System and Method for the Synthesis of an Economic Web and the Identification of New Market Niches*.

FIG. 5 discloses a representative computer system 510 in conjunction with which the embodiments of the present invention may be implemented. Computer system 510 may be a personal computer, workstation, or a larger system such as a minicomputer. However, one skilled in the art of computer systems will understand that the present invention is not limited to a particular class or model of computer.

As shown in FIG. 5,representative computer system 510 includes a central processing unit (CPU) 512, a memory unit 514, one or more storage devices 516, an input device 518, an output device 520, and communication interface 2922. A system bus 524 is provided for communications between these elements. Computer system 510 may additionally function through use of an operating system such as Windows, DOS, or UNIX. However, one skilled in the art of computer systems will understand that the present invention is not limited to a particular configuration or operating system.

Storage devices 516 may illustratively include one or more floppy or hard disk drives, CD-ROMs, DVDs, or tapes. Input device 518 comprises a keyboard, mouse, microphone, or other similar device. Output device 520 is a computer monitor or any other known computer output device. Communication interface 522 may be a modem, a network interface, or other connection to external electronic devices, such as a serial or parallel port

While the above invention has been described with reference to certain preferred embodiments, the scope of the present invention is not limited to these embodiments. One skill in the art may find variations of these preferred embodiments which, nevertheless, fall within the spirit of the present invention, whose scope is defined by the claims set forth below.

Claims

Claims

1. A method for partitioning a space of data comprising the steps of:

choosing a plurality of dissimilarity measures;

partitioning the space for each of said plurality of dissimilarity measures;

evaluating said partitioning for each of said plurality of dissimilarity measures; and

selecting one or more of said dissimilarity measures on the basis of said evaluation.

2. A method for partitioning a space as in claim 1 further comprising the steps of:

performing at least one operation of said selected dissimilarity measures to generate a new plurality of dissimilarity measures; and

repeating said partitioning the space step and said evaluating said partitioning step for each of said new plurality of dissimilarity measures; and

selecting one or more of said new dissimilarity measures on the basis of said evaluation.

3. A method for partitioning a space as in claim 2 further comprising the step of iterating on said performing at least one operation of said selected dissimilarity measures step, said repeating said partitioning the space step and said selecting one or more of said new dissimilarity measures step to achieve an optimal partition.

4. A method for partitioning a space as in claim 2 wherein said at least one operation is a genetic operation.

5. A method for partitioning a space as in claim 4 wherein said genetic operation is selected from the group consisting of a mutation and a crossover.

6. A method for partitioning a space as in claim 1 wherein said choosing a plurality of dissimilarity measures step comprises the steps of:

choosing a family of dissimilarity measures; and

randomly generating said plurality of dissimilarity measures from said chosen family.

7. A method for partitioning a space as in claim 1 wherein said partitioning for each of said plurality of dissimilarity measures step is performed by at least one clustering algorithm.

8. A method for partitioning a space as in claim 7 wherein said at least one clustering algorithm is a k-means clustering algorithm.

9. A method for partitioning a space as in claim 6 wherein said dissimilarity measure is a dissimilarity function and said family of dissimilarity measures is a family of dissimilarity functions.

10. A method for partitioning a space as in claim 9 wherein said family of dissimilarity functions for said space having two dimensions is:

$$d_{tu} = \left| f_1(x_{i1}, x_{i2}, x_{u1}, x_{u2})(x_{i1} - x_{u1})^2 + f_2(x_{i1}, x_{i2}, x_{u1}, x_{u2})(x_{i2} - x_{u2})^2 \right|^{1/2},$$

wherein

x_{i1}, x_{i2}, x_{u1}, and x_{u2} are said data in said space;

f_1 and f_2 are polynomial functions of their variables; and

f_1 and f_2 are characterized by a plurality of parameters.

11. A method for partitioning a space as in claim 10 wherein said polynomial functions are second degree polynomial functions.

12. A method for partitioning a space as in claim 10 wherein said plurality of parameters are coefficients of the polynomials.

13. A method for partitioning a space as in claim 12 wherein the variations of said plurality of parameters are restricted to [-10,10].

14. A method for partitioning a space as in claim 12 wherein said evaluating said partitioning for each of said plurality of dissimilarity measures step comprises the step of assigning a fitness to said each of said plurality of dissimilarity measures and said fitness is defined by:

$$F^{v} = \frac{1}{1 + M_{in} + M_{out}},$$

wherein:
M_{in} is the number of said data that are assigned to the same partition that do not belong in the same partition; and
M_{out} is the number of said data that are assigned to different partitions and do belong in the same partition.

15. A method for partitioning a space as in claim 6 wherein said family of dissimilarity measures are general function approximators.

16. A method for partitioning a space as in claim 6 wherein said general function approximators are neural networks having connections weights.

17. A method for partitioning a space as in claim 1 wherein said evaluating said partitioning for each of said plurality of dissimilarity measures step is performed by a human observer.

18. A method for partitioning a space as in claim 17 wherein said selecting one or more of said dissimilarity measures on the basis of said evaluation step is performed by a human observer based on said evaluation.

19. Computer executable software code stored on a computer readable medium, the code for partitioning a space of data, the code comprising:

code to choosing a plurality of dissimilarity measures;

code to partition the space for each of said plurality of dissimilarity measures;

code to evaluate said partitioning for each of said plurality of dissimilarity measures; and

code to select one or more of said dissimilarity measures on the basis of said evaluation.

20. Computer executable software code stored on a computer readable medium, the code for partitioning a space as in claim 19, the code further comprising:

code to perform at least one operation of said selected dissimilarity measures to generate a new plurality of dissimilarity measures;

code to repeat said partitioning the space step and said evaluating said partitioning step for each of said new plurality of dissimilarity measures; and

code to select one or more of said new dissimilarity measures on the basis of said evaluation.

21. A programmed computer system for partitioning a space comprising at least one memory having at least one region storing computer executable program code and at least one processor for executing the program code stored in said memory, wherein the program code includes:

code to choosing a plurality of dissimilarity measures;

code to partition the space for each of said plurality of dissimilarity measures;

code to evaluate said partitioning for each of said plurality of dissimilarity measures; and

code to select one or more of said dissimilarity measures on the basis of said evaluation.

22. A programmed computer system for partitioning a space comprising at least one memory having at least one region storing computer executable program code and at least one processor for executing the program code stored in said memory as in claim 21, wherein the program code further includes:

code to perform at least one operation of said selected dissimilarity measures to generate a new plurality of dissimilarity measures;

code to repeat said partitioning the space step and said evaluating said partitioning step for each of said new plurality of dissimilarity measures; and

code to select one or more of said new dissimilarity measures on the basis of said evaluation.

23. A method for determining customer demand for products comprising the steps of:

defining a space having R dimensions wherein each point in said space corresponds to a vector of properties;

constructing a landscape for said space comprising the steps of:

locating at least one point on said space where a predetermined customer would purchase a product having said corresponding vector of properties at a predetermined price; and

sampling a set of points on an R-dimensional sphere surrounding said selected point at a predetermined step length from said selected point to determine a first subset of said set of points where the predetermined customer would make a purchase at said predetermined price and to

determine a second subset of said sampled points where the customers would not make the purchase at said predetermined price, said first subset of points and said second subset of points form at least one indifference surface between a buying region and a non-buying region at said predetermined price.

24. A method for determining customer demand for products as in claim 23 wherein said constructing a landscape for said space further comprises the steps of:

selecting at least one point on said indifference surface; and

repeating said sampling step from said selected point to extend said at least one indifference surface.

25. A method for determining customer demand for products as in claim 24 further comprising the step of iterating on said selecting at least one point on said indifference surface step and said repeating said sampling step from said selected point step to further extend said indifference surface.

26. A method for determining customer demand for products as in claim 23 further comprising the steps of determining characteristics of said indifference surface from said sampling step.

27. A method for determining customer demand for products as in claim 26 wherein said indifference surface characteristics comprise a degree of ruggedness.

28. A method for determining customer demand for products as in claim 26 wherein said indifference surface characteristics comprise at least one correlation length.

29. A method for determining customer demand for products as in claim 26 further comprising the step of

locating one or more points on said indifference surface having a small amount of said corresponding vector of properties to identify peaks on said indifference surface.

30. A method for determining customer demand for products as in claim 29 wherein said indifference surface characteristics further comprise at least one typical walk length to said identified peaks.

31. A method for determining customer demand for products as in claim 29 wherein said indifference surface characteristics further comprise at least one clustering measure of said identified peaks.

32. A method for determining customer demand for products as in claim 29 further comprising the steps of:

defining a family of possible models to represent the customer demand; and

selecting one or more models from said family of possible models that are compatible with said indifference surface characteristics.

33. A method for determining customer demand for products as in claim 32 wherein said selected models have a plurality of parameters.

34. A method for determining customer demand for products as in claim 33 further comprising the step of determining values of said plurality of parameters for said selected models from said sampling step.

35. A method for determining customer demand for products as in claim 33 wherein said values of said plurality of selected parameters are determined using Bayesian analysis.

36. Computer executable software code stored on a computer readable medium, the code for determining customer demand for products, the code comprising:

code to define a space having R dimensions wherein each point in said space corresponds to a vector of properties;

code to construct a landscape for said space comprising:

code to locate at least one point on said space where a predetermined customer would purchase a product having said corresponding vector of properties at a predetermined price; and

code to sample a set of points on an R-dimensional sphere surrounding said selected point at a predetermined step length from said selected point to determine a first subset of said set of points where the predetermined customer would make a purchase at said predetermined price and to determine a second subset of said sampled points where the customers would not make the purchase at said predetermined price, said first subset of points and said second subset of points form at least one indifference surface between a buying region and a non-buying region at said predetermined price.

37. Computer executable software code stored on a computer readable medium, the code for determining customer demand for products as in claim 36, wherein said code to construct a landscape for said space further comprises:

code to select at least one point on said indifference surface; and

code to repeat said sampling step from said selected point to extend said at least one indifference surface.

38. Computer executable software code stored on a computer readable medium, the code for determining customer

demand for products as in claim 37, the code further comprising:

code to iterate on said selecting at least one point on said indifference surface step and said repeating said sampling step from said selected point step to further extend said indifference surface.

39. A programmed computer system for determining customer demand for products comprising at least one memory having at least one region storing computer executable program code and at least one processor for executing the program code stored in said memory, wherein the program code includes:

code to define a space having R dimensions wherein each point in said space corresponds to a vector of properties;

code to construct a landscape for said space comprising:

code to locate at least one point on said space where a predetermined customer would purchase a product having said corresponding vector of properties at a predetermined price; and

code to sample a set of points on an R-dimensional sphere surrounding said selected point at a predetermined step length from said selected point to determine a first subset of said set of points where the predetermined customer would make a purchase at said predetermined price and to determine a second subset of said sampled points where the customers would not make the purchase at said predetermined price, said first subset of points and said second subset of points form at least one indifference surface between a buying region and a non-buying region at said predetermined price.

40. A programmed computer system for determining customer demand for products comprising at least one memory having at least one region storing computer executable

program code and at least one processor for executing the program code stored in said memory as in claim 39, wherein said code to construct a landscape for said space further comprises:

code to select at least one point on said indifference surface; and

code to repeat said sampling step from said selected point to extend said at least one indifference surface.

41. A programmed computer system for determining customer demand for products comprising at least one memory having at least one region storing computer executable program code and at least one processor for executing the program code stored in said memory as in claim 40, wherein said code further comprises:

code to iterate on said selecting at least one point on said indifference surface step and said repeating said sampling step from said selected point step to further extend said indifference surface.

42. A method for creating a model of consumer preferences from consumer data comprising the steps of:

constructing a plurality of candidate maps form the consumer data to actual consumer preferences;

constructing a family of agent-based models;

evaluating said plurality of candidate maps and said family of agent-based models with respect to said consumer data;

selecting one or more of said plurality of candidate maps and said family of agent based models based on said evaluation; and

performing at least one operation on said selected candidate maps and said selected agent-based models to generate a new plurality of candidate maps and a new family of agent-based models.

43. A method for creating a model of consumer preferences from consumer data as in claim 42 further comprising the step of iterating on said evaluating said plurality of candidate maps step, said selecting one or more of said plurality of candidate maps and said family of agent based models step and said performing at least one operation on said selected candidate maps and said selected agent-based models step to achieve an optimal model of consumer preferences.

44. A method for creating a model of consumer preferences from consumer data as in claim 42 wherein said at least one operation is a genetic operation.

45. A method for creating a model of consumer preferences from consumer data as in claim 44 wherein said genetic operation is selected from the group consisting of a mutation and a crossover.

46. Computer executable software code stored on a computer readable medium, the code for creating a model of consumer preferences from consumer data, the code comprising:

code to construct a plurality of candidate maps form the consumer data to actual consumer preferences;

code to construct a family of agent-based models;

code to evaluate said plurality of candidate maps and said family of agent-based models with respect to said consumer data;

code to select one or more of said plurality of candidate maps and said family of agent based models based on said evaluation; and

code to perform at least one operation on said selected candidates maps and said selected agent-based models to generate a new plurality of candidate maps and a new family of agent-based models.

47. A programmed computer system for creating a model of consumer preferences from consumer data, comprising at least one memory having at least one region storing computer executable program code and at least one processor for executing the program code stored in said memory, wherein the program code includes:

code to construct a plurality of candidate maps from the consumer data to actual consumer preferences;

code to construct a family of agent-based models;

code to evaluate said plurality of candidate maps and said family of agent-based models with respect to said consumer data;

code to select one or more of said plurality of candidate maps and said family of agent based models based on said evaluation; and

code to perform at least one operation on said selected candidate maps and said selected agent-based models to generate a new plurality of candidate maps and a new family of agent-based models.

48. A method for marketing and introducing novel products from consumer data comprising the steps of:

creating a model of customer preferences; and

identifying novel products using the method for determining customer demand of claim 23.

49. A method for marketing and introducing novel products from consumer data wherein said creating a model of customer preferences step is performing using the method of claim 1.

50. Computer executable software code stored on a computer readable medium, the code for marketing and introducing novel products from consumer data, the code comprising:

code to create a model of customer preferences; and

code to identify novel products using the method for determining customer demand of claim 23.

51. Computer executable software code stored on a computer readable medium, the code for marketing and introducing novel products from consumer data as in claim 50, wherein said code to create a model of customer preferences is the code of claim 19.

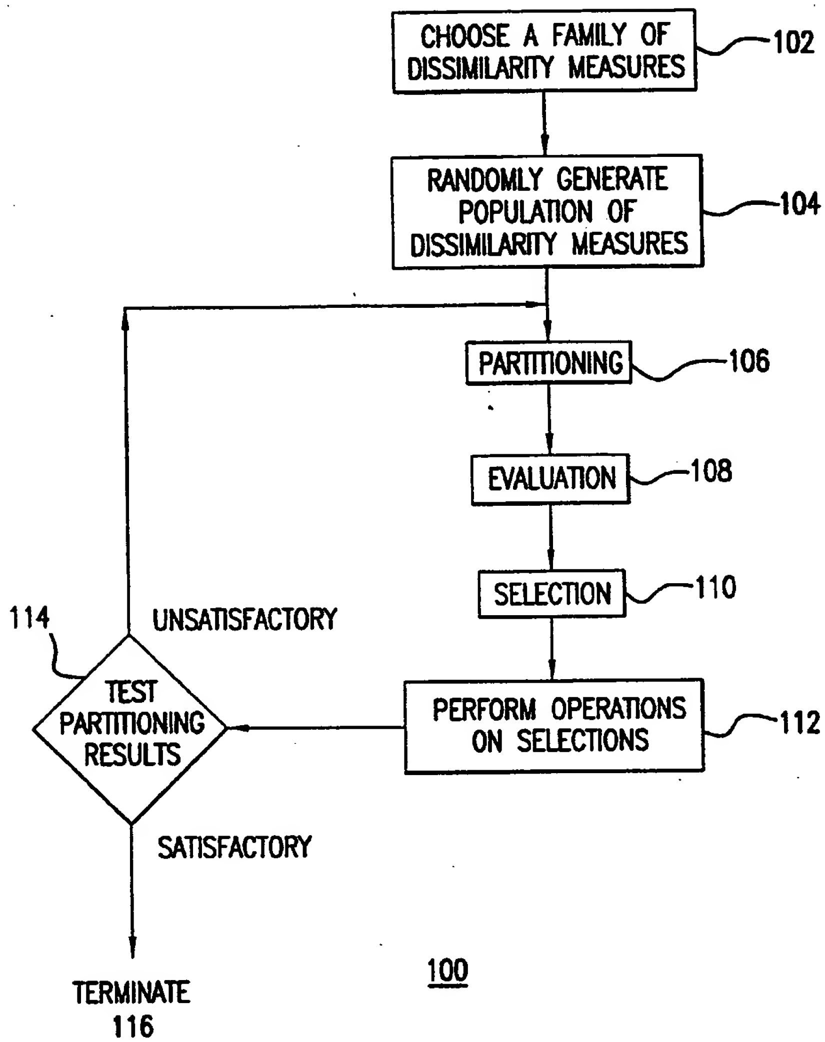

CHOOSE A FAMILY OF DISSIMILARITY MEASURES
102
RANDOMLY GENERATE POPULATION OF DISSIMILARITY MEASURES
104
PARTITIONING
106
EVALUATION
108
SELECTION
110
PERFORM OPERATIONS ON SELECTIONS
112
114
TEST PARTITIONING RESULTS
UNSATISFACTORY
SATISFACTORY
TERMINATE
116
100

FIG.1

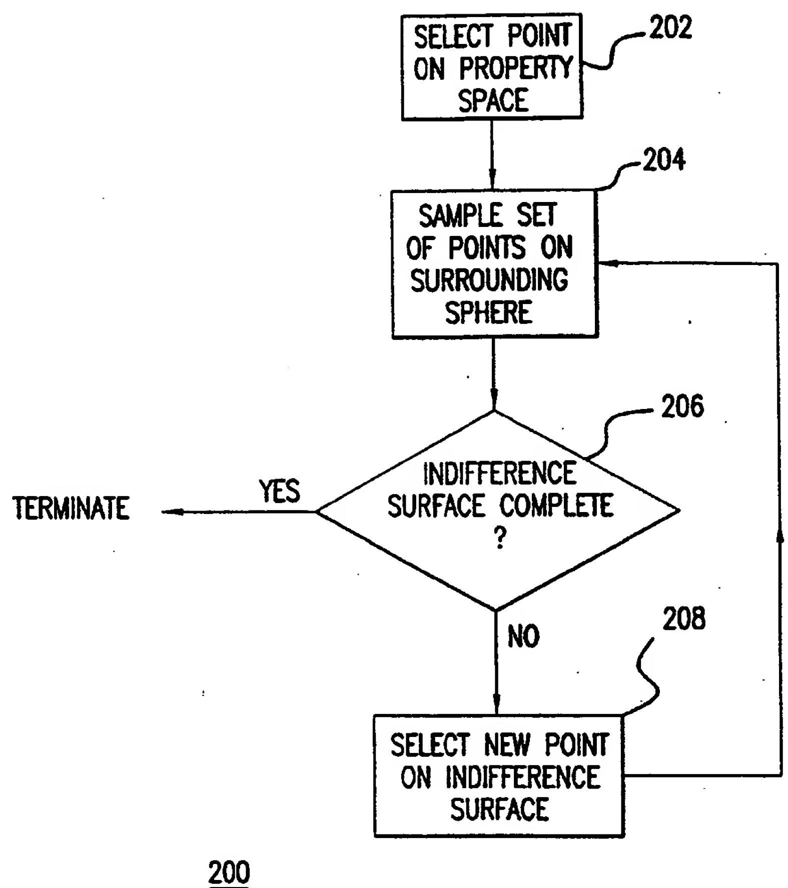

SELECT POINT ON PROPERTY SPACE
202
SAMPLE SET OF POINTS ON SURROUNDING SPHERE
204
INDIFFERENCE SURFACE COMPLETE ?
206
YES
TERMINATE
NO
SELECT NEW POINT ON INDIFFERENCE SURFACE
208
200

FIG.2

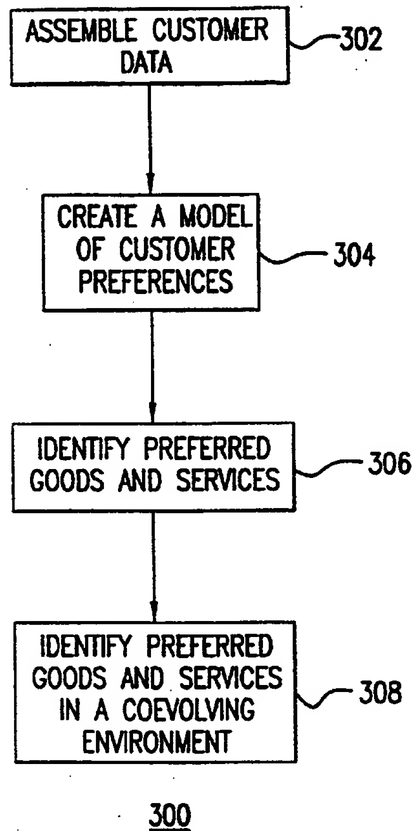

ASSEMBLE CUSTOMER DATA
302
CREATE A MODEL OF CUSTOMER PREFERENCES
304
IDENTIFY PREFERRED GOODS AND SERVICES
306
IDENTIFY PREFERRED GOODS AND SERVICES IN A COEVOLVING ENVIRONMENT
308
300

FIG.3

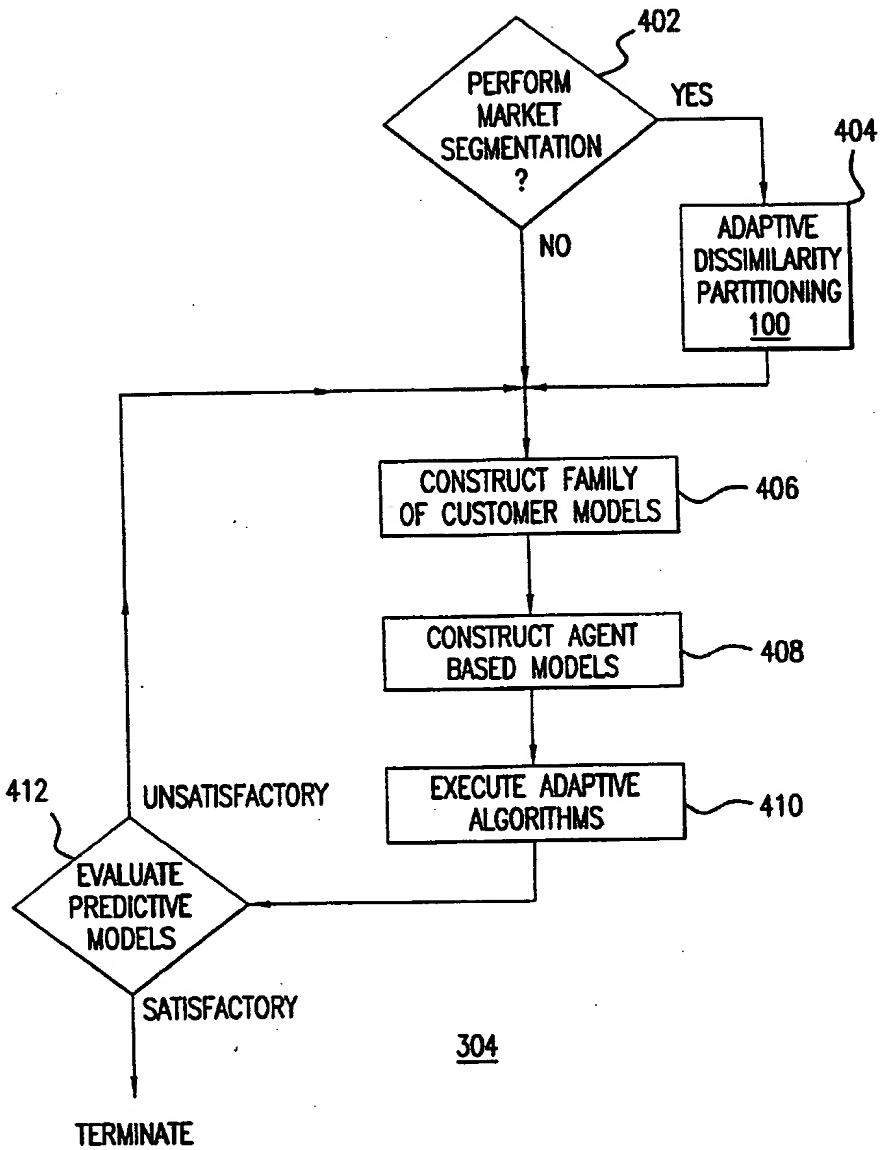

PERFORM MARKET SEGMENTATION ?
402
YES
NO
404
ADAPTIVE DISSIMILARITY PARTITIONING 100
CONSTRUCT FAMILY OF CUSTOMER MODELS
406
CONSTRUCT AGENT BASED MODELS
408
EXECUTE ADAPTIVE ALGORITHMS
410
412
UNSATISFACTORY
EVALUATE PREDICTIVE MODELS
SATISFACTORY
TERMINATE
304

FIG.4

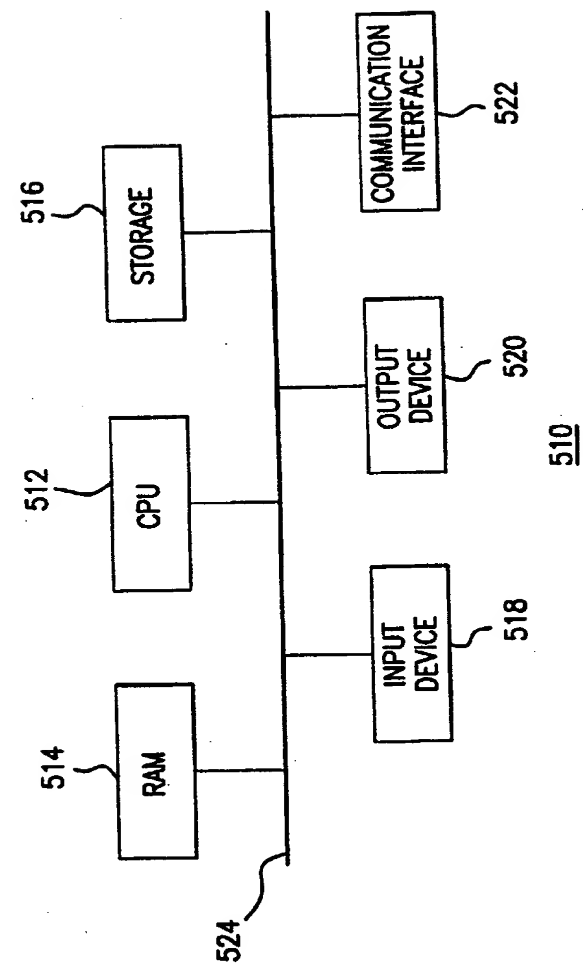

STORAGE
516
CPU
512
RAM
514
524
COMMUNICATION INTERFACE
522
OUTPUT DEVICE
520
INPUT DEVICE
518
510

FIG.5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/15236

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :G06F 15/21
US CL :705/10; 706/19
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 705/7, 8, 9, 10, 11, 37; 706/13, 19, 62

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
APS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A, P	US 5,808,908 A (GHAHRAMANI) 15 September 1998, all.	1-48, 50
A, P	US 5,796,924 A (ERRICO et al) 18 August 1998, all.	1-48, 50
A	US 5,724,262 A (GHAHRAMANI) 03 March 1998, all.	1-48, 50
A	US 5,461,698 A (SCHWANKE et al) 24 October 1995, all.	1-48, 50
A	US 5,041,972 A (FROST) 20 August 1991, all.	1-48, 50
A	US 4,529,228 A (KRAMER) 17 July 1985, all.	1-48, 50

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
07 SEPTEMBER 1999	27 OCT 1999

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	ALLEN MACDONALD James R. Matthews Telephone No. (703) 305-9708

Form PCT/ISA/210 (second sheet)(July 1992)★

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [ ] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. [X] Claims Nos.: 49
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Claim 49 embodied two of the three groups of restricted inventions; therefore, the search for this single claim would have encompassed a multitude of classes/subs.

3. [X] Claims Nos.: 51
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. [X] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[X] No protest accompanied the payment of additional search fees.

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-22, drawn to partitioning a space of data.
Group II, claims 23-41, 48, and 50, drawn to determining customer demand for products.
Group III, claims 42-47, drawn to creating a model of consumer preferences from consumer data.

The inventions listed as Groups I, II, and III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The three inventions are sub-combinations which are usable together, but may also function independently of each other.